COMMUNICATION SCIENCES AND ENGINEERING

XIII. STATISTICAL COMMUNICATION THEORY*

Prof. Y. W. Lee
Prof. A. G. Bose
Prof. I. M. Jacobs
Prof. D. J. Sakrison
Prof. M. Schetzen
Prof. H. L. Van Trees, Jr.
V. R. Algazi
A. A. Al-Shalchi
R. Alter
D. S. Arnstein
M. E. Austin
R. F. Bauer
J. D. Bruce
A. M. Bush
D. F. DeLong, Jr.
A. G. Gann
C. E. Gray
H. Hemami
W. F. Kelly
V. S. Mummert
R. B. Parente
C. E. Wernlein, Jr.
W. S. Widnall
R. W. Witte

A. WORK COMPLETED

1. DESIGN OF A WIENER-LEE VARIABLE FILTER

This study has been completed by H. Hemami. In January 1962 he submitted the results to the Department of Electrical Engineering, M.I.T., as a thesis in partial fulfillment of the requirements for the degree of Master of Science.

A. G. Bose

2. MEASUREMENT OF A SECOND-DEGREE WIENER KERNEL IN A NONLINEAR SYSTEM BY CROSSCORRELATION

The present study has been completed by W. S. Widnall. It was submitted as a thesis in partial fulfillment of the requirements for the degree of Master of Science, Department of Electrical Engineering, M.I.T., January 1962.

Y. W. Lee

B. A PROPERTY OF OPTIMUM SYSTEMS

In this report, we consider an important property of systems for the separation of signals. In Fig. XIII-1, $f_r(t)$ is the sum of two signals, $f_1(t)$ and $f_2(t)$, so that

$$f_r(t) = f_1(t) + f_2(t). \tag{1}$$

The signals are not restricted to any particular type of functions. They can be periodic, aperiodic or random; they could be dependent or independent of each other. In Fig. XII-1 system A has been designed on the basis of some error criterion for the desired output $f_1(t)$. Similarly, system B has been designed on the basis of the same error criterion for the desired output $f_2(t)$. The error criterion can be any function of the magnitude of the error.

We shall show that, if the network A is the optimum N^{th}-order nonlinear network for which some function of the magnitude of the error, $E_1(t)$, between the

*This work was supported in part by the National Institutes of Health (Grant MH-04737-02).

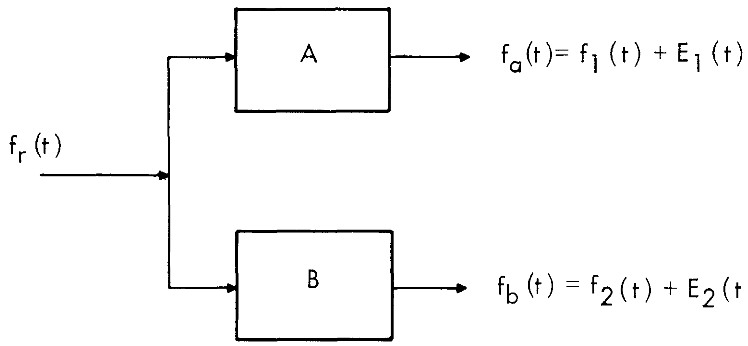


Fig. XIII-1. Optimum systems A and B.

actual output, $f_a(t)$, and the desired output, $f_1(t)$, is a minimum, then, for the same error criterion, the optimum N^{th}-order nonlinear system B has the output

$$f_b(t) = f_r(t) - f_a(t) \tag{2}$$

and thus can be synthesized as shown in Fig. XIII-2.

To show this relationship, we note, first, that if the system A is an N^{th}-order nonlinear system, then the system depicted in Fig. XIII-2 is also an N^{th}-order nonlinear system. We note, then, that by substituting Eq. 1 and the relation

$$f_a(t) = f_1(t) + E_1(t) \tag{3}$$

in Eq. 2, we obtain

$$f_b(t) = f_2(t) - E_1(t). \tag{4}$$

That is, the error between the actual output of the system of Fig. XIII-2 and the desired output, $f_2(t)$, is the negative of $E_1(t)$, the error of system A. This result implies that

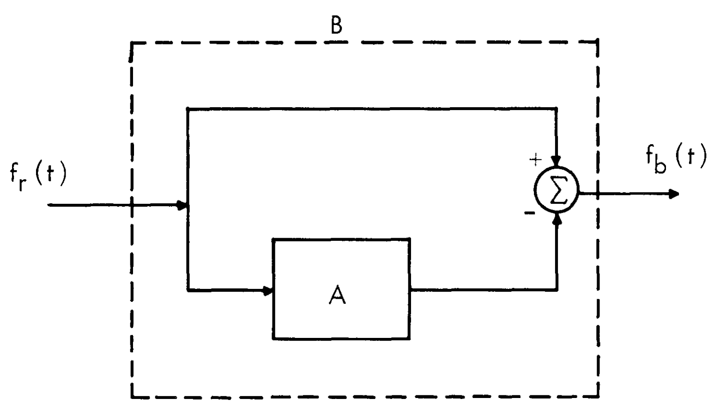


Fig. XIII-2. Form of optimum system B.

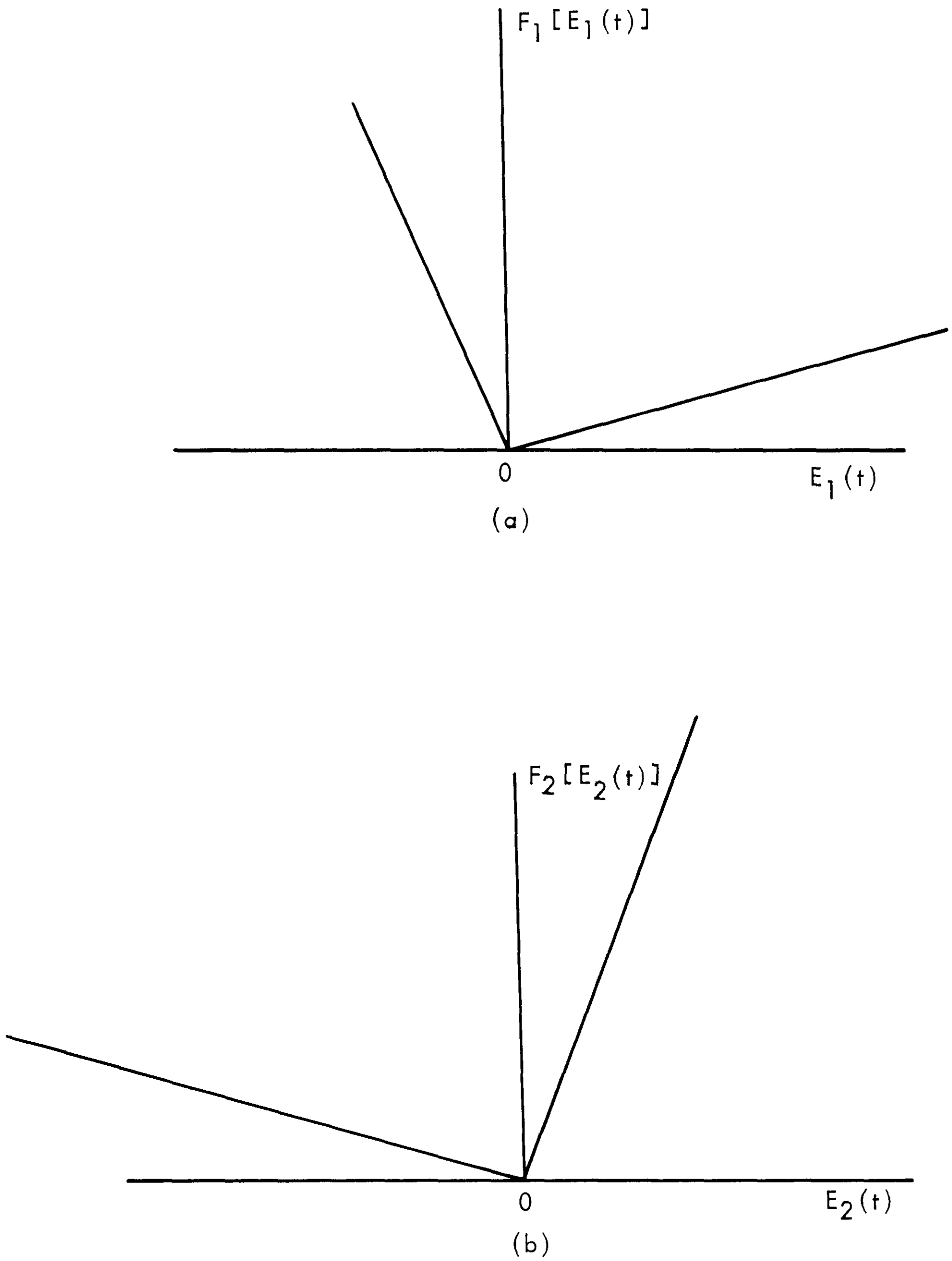


Fig. XIII-3. (a) Example of a function of the error for system A. (b) Corresponding function of the error for system B illustrated by Fig. XIII-2.

the system of Fig. XIII-2 is the optimum N^{th}-order nonlinear system for the desired output, $f_2(t)$. For, assume that it is not; then there is another system C, with the output $f_c(t)$ for which the function of the magnitude of the error, $E_3(t) = f_c(t) - f_2(t)$, is smaller than that of the system of Fig. XIII-2. We could then construct the N^{th}-order nonlinear system D with the output $f_r(t) - f_c(t) = f_1(t) - E_3(t)$. For the desired output $f_1(t)$, the function of the magnitude of the error for system D would then be smaller than that of the system A. This is a contradiction, since we have assumed that system A is the optimum N^{th}-order nonlinear system for the desired output $f_1(t)$. Thus we conclude that the system shown in Fig. XIII-2 is the optimum N^{th}-order system for the desired output $f_2(t)$.

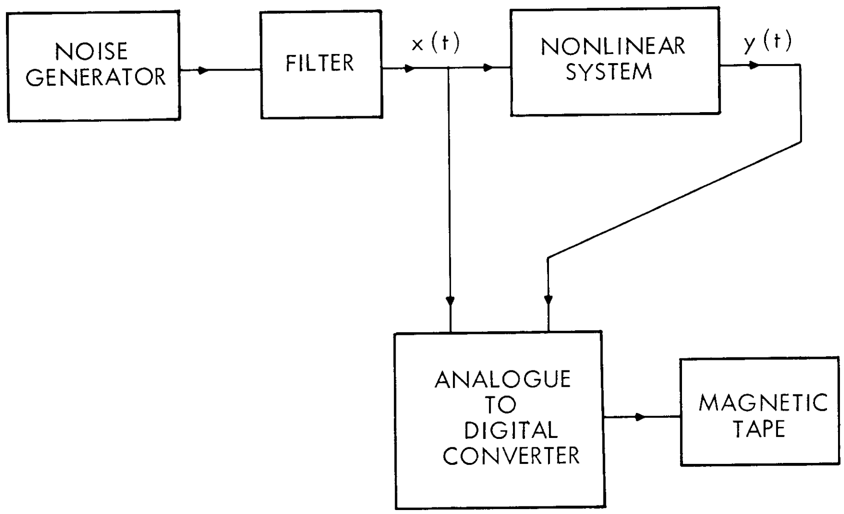
Note that this result does not depend in any manner upon the statistical dependence between $f_1(t)$ and $f_2(t)$ or upon the form of the system A. Also, the result does not depend in any manner upon the criterion of the magnitude of the error. In whatever sense system A is optimum for the desired output $f_1(t)$, system B, shown in Fig. XIII-2, is also optimum for the desired output $f_2(t)$. Furthermore, since the magnitude of the error for system A is equal to that of system B, any function of the magnitude of the error will be the same for both systems.

This result can be generalized. We note from Eq. 4 that the error for system B illustrated by Fig. XIII-2 is the negative of the error for system A. By choosing the error criterion for the design of system A to be any function of the magnitude of the error, we obviated this difference, since the error criterion was then a symmetric function of the error. If, however, the error criterion were not a symmetric function of the error, then system B illustrated by Fig. XIII-2 would be optimum for an error criterion that is the mirror image of that used for the design of system A. Thus if system A were designed to minimize the probability that the error is positive, then system B would minimize the probability that the error is negative. As another example, if the error criterion for system A is the average of the function of the error shown in Fig. XIII-3a, then system B is optimum for an error criterion that is the average of the function of the error shown in Fig. XIII-3b.

M. Schetzen

C. MEASUREMENT OF A SECOND-DEGREE WIENER KERNEL IN A NONLINEAR SYSTEM BY CROSSCORRELATION

Wiener has developed a functional representation for nonlinear systems.[1] Lee and Schetzen have shown theoretically that the kernels of the functional representation of an unknown system may be measured by a crosscorrelation of the output of the nonlinear system with a multidimensional product formed from the input to the system when the input is broadband Gaussian noise.[2] The work reported here is the first experimental example of the Lee-Schetzen method of measuring the kernels.[3]



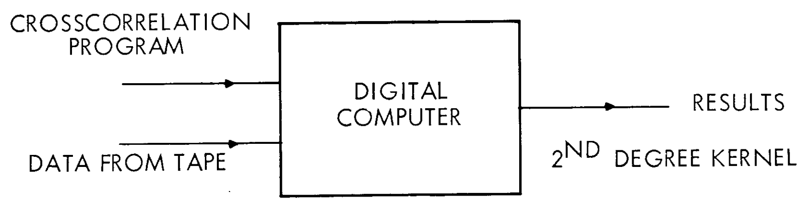


Fig. XIII-4. Block diagram of the experiment.

A simple nonlinear system was chosen – a linear RLC circuit whose response is squared by a memoryless squarer. The system was driven by an appropriate filtered noise source. The input and output waveforms were sampled, converted to binary digital numbers, and written on magnetic tape. A digital computer calculated the second-degree crosscorrelation from the sampled data on tape. The results show clearly the second-degree Wiener kernel of this nonlinear system, and establish the practicality of this method of kernel measurement. The experiment is summarized in Fig. XIII-4.

1. Theory of Kernel Measurement

The output y(t) of a nonlinear system is functionally related to the system input x(t). Wiener has shown that this functional relationship may be expanded in a functional infinite series

$$y(t) = \sum_{n=0}^{\infty} G_n[g_n, x(t)].$$

The first four terms in this functional expansion are

$$G_o[g_o, x(t)] = g_o$$

$$G_1[g_1, x(t)] = \int_{-\infty}^{\infty} g_1(\tau_1)\, x(t-\tau_1)\, d\tau_1$$

$$G_2[g_2, x(t)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_2(\tau_1, \tau_2)\, x(t-\tau_1)\, x(t-\tau_2)\, d\tau_1 d\tau_2 - P\int_{-\infty}^{\infty} g_2(\tau_1, \tau_1)\, d\tau_1$$

$$G_3[g_3, x(t)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_3(\tau_1, \tau_2, \tau_3)\, x(t-\tau_1)\, x(t-\tau_2)\, x(t-\tau_3)\, d\tau_1 d\tau_2 d\tau_3$$

$$- 3P\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_3(\tau_1, \tau_2, \tau_2)\, x(t-\tau_1)\, d\tau_1 d\tau_2.$$

Any nonlinear system that is not time-variant and whose present output cannot depend on the infinite past has such a G-functional representation. The kernels $g_n(\tau_1, \tau_2, \ldots \tau_n)$ are properties of the system alone and do not depend upon the nature of the input x(t). Once one has determined the kernels of a system, then the system is completely characterized; the output may be computed for any given input.

Lee and Schetzen have shown theoretically that the kernels of an unknown system can be measured by crosscorrelation. In particular, the n^{th}-degree system kernel is proportional to the crosscorrelation of the system output with an N-dimensional product formed from the input, provided that the input is broadband Gaussian noise.

$$g_n(\sigma_1, \sigma_2, \ldots \sigma_n) = \frac{1}{N!\, P^n}\, \overline{y(t)\, x(t-\sigma_1)\, x(t-\sigma_2)\, \ldots\, x(t-\sigma_n)}^{\text{time average}}$$

where all of the delays σ_i must be distinct. The constant P is the power density of the broadband noise at the frequency band that excites the system. P is measured in input units squared per cycle per second.

2. Experimental Procedure

The particular electrical system chosen for measurement was a linear RLC network

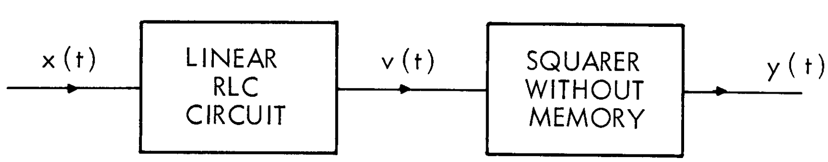


Fig. XIII-5. The particular nonlinear system that was measured.

followed by an electronic squarer, as shown in Fig. XIII-5. Mathematically, such a system has a simple functional representation. The output v(t) of the linear network is given by

$$v(t) = \int_{-\infty}^{\infty} h(\tau)\, x(t-\tau)\, d\tau,$$

where h(t) is the unit impulse response of the linear network. The output of the squarer is

$$y(t) = K[v(t)]^2$$

$$= K\left[\int_{-\infty}^{\infty} h(\tau)\, x(t-\tau)\, d\tau\right]^2$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Kh(\tau_1)\, h(\tau_2)\, x(t-\tau_1)\, x(t-\tau_2)\, d\tau_1 d\tau_2.$$

This functional relationship can be expanded in terms of G-functionals. The resulting expansion has only two terms — a zero-degree term and a second-degree term

$$y(t) = G_o[g_o, x(t)] + G_2[g_2, x(t)],$$

in which

$$G_o = P\int_{-\infty}^{\infty} Kh^2(\tau_1)\, d\tau_1$$

$$G_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Kh(\tau_1)\, h(\tau_2)\, x(t-\tau_1)\, x(t-\tau_2)\, d\tau_1 d\tau_2 - P\int_{-\infty}^{\infty} Kh^2(\tau_1)\, d\tau_1.$$

The second-degree kernel is

$$g_2(\tau_1, \tau_2) = Kh(\tau_1)\, h(\tau_2).$$

The input and output of the linear part of the system are voltages; thus the unit impulse response has the dimensions of inverse seconds. The actual response of the RLC network to a pulse of unit area is shown in Fig. XIII-6. The input-output characteristic of the squarer is shown in Fig. XIII-7.

The nonlinear system was driven by a broadband Gaussian noise source. The output of a shot-noise generator was passed through an adjustable cutoff lowpass filter. The bandwidth of the resulting noise was chosen to be broad compared with the band of frequencies that excite the system, but not so broad as to cause a large variance of the crosscorrelation calculations. The power density spectrum of the noise generator-filter combination is shown in Fig. XIII-8.

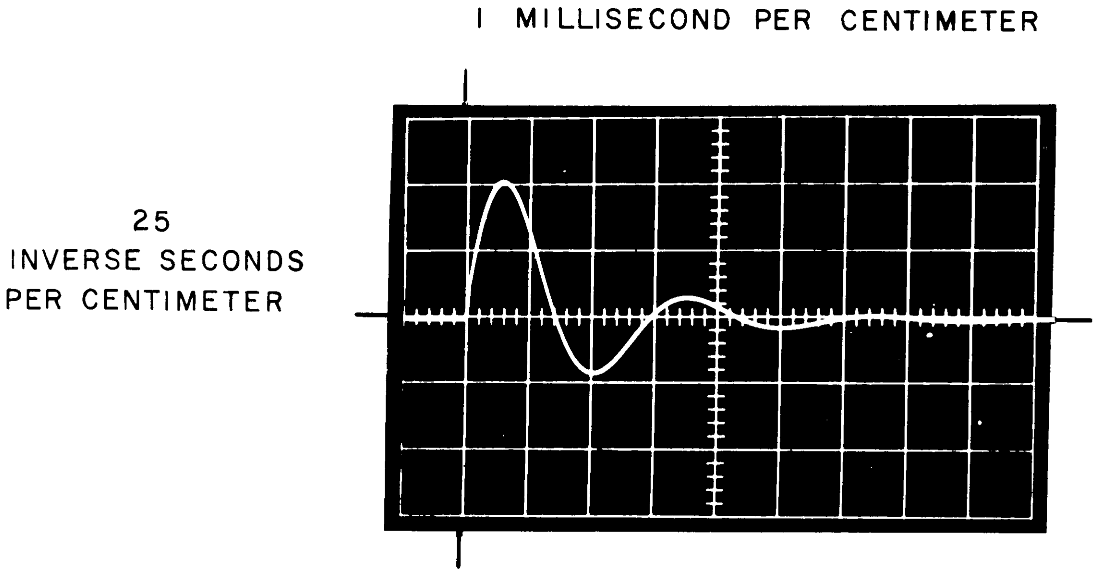


Fig. XIII-6. Impulse response of the linear RLC network.

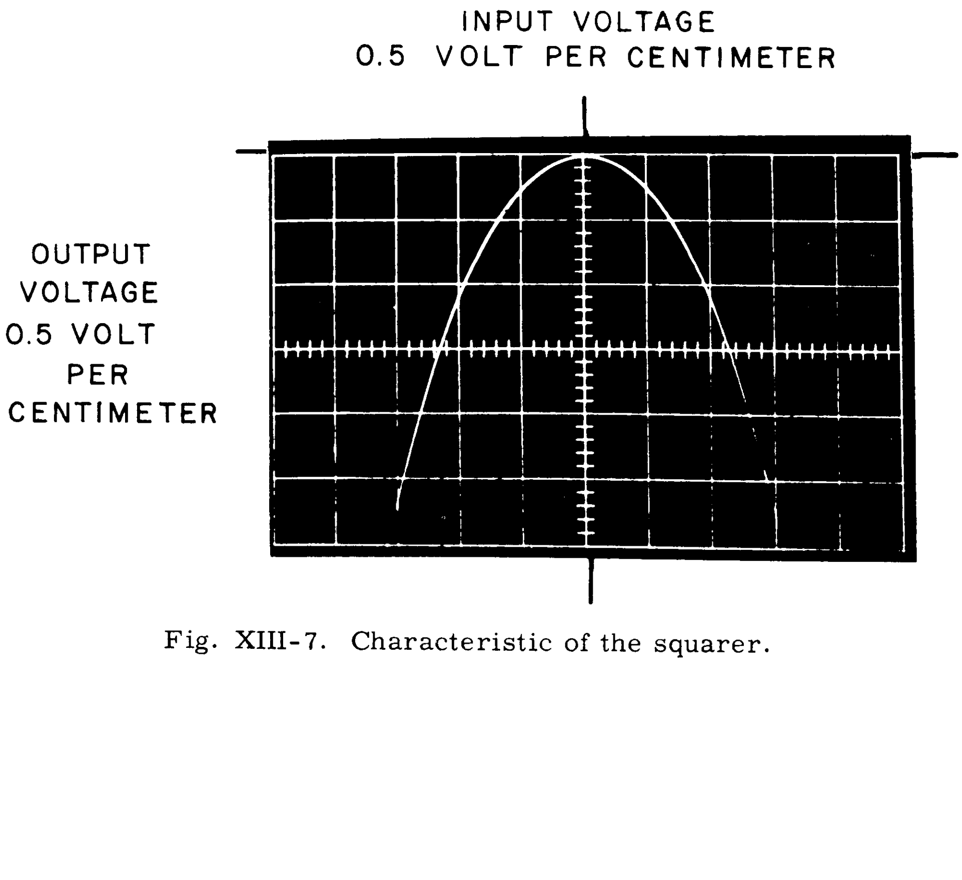


Fig. XIII-7. Characteristic of the squarer.

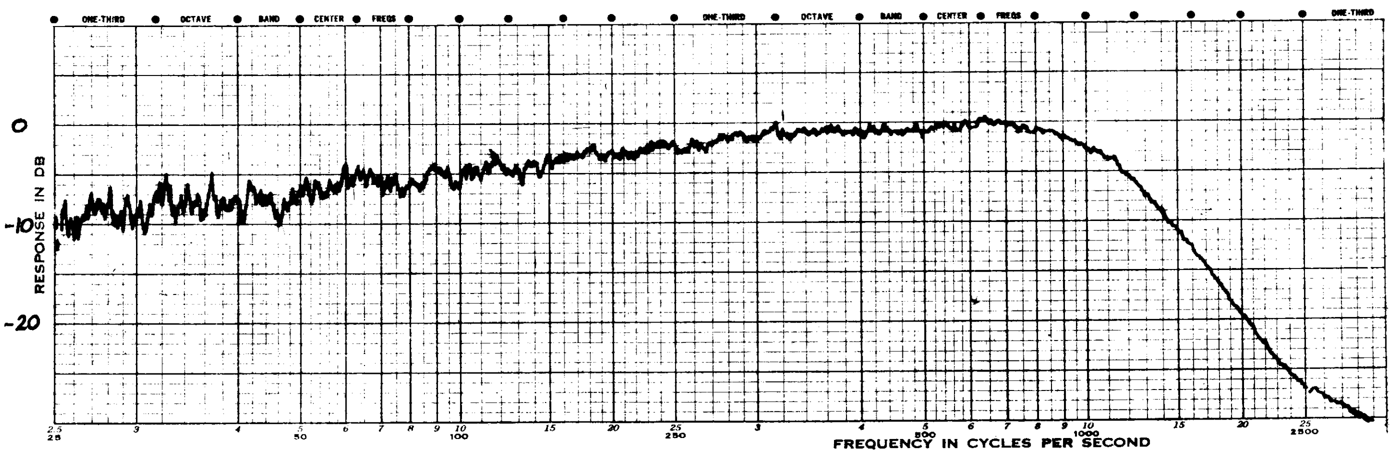


Fig. XIII-8. Power density spectrum of the filtered noise source.

The data needed for the crosscorrelation were recorded automatically by a dual-channel analog-to-digital conversion system. The system sampled simultaneously the input and output voltages of the nonlinear system. Every 0.269 msec a new sample pair was measured. The voltages were converted to 11-bit binary numbers, and these were recorded on magnetic tape.

A suitable second-degree crosscorrelation program was written for the IBM 7090 digital computer. The program calculates the second-degree Wiener kernel by using the Lee-Schetzen theorem

$$g_2(\sigma_1, \sigma_2) = \frac{1}{2P^2} \overline{y(t)\, x(t-\sigma_1)\, x(t-\sigma_2)}.$$

The kernel cannot be calculated for continuous values of input delay because the data were sampled at discrete intervals. The kernel is calculated as a discrete array whose elements are the values of the kernel at arguments equal to integral multiples of the basic sampling interval.

Notice that the value of the kernel is unchanged if σ_1 and σ_2 are interchanged. This property of kernel symmetry is used to save computer time. Only one-half of the elements in the kernel array need be directly calculated by crosscorrelation. The other symmetric one-half are set to equal values at the end of the program.

3. Crosscorrelation Results

The computer results have been used to make a three-dimensional model of the second-degree kernel. This is shown in Fig. XIII-9. A few cross sections of the kernel with the numerical values given are shown in Fig. XIII-10. In both the model and the graphs, the kernel has been inverted so that the greatest peak is not shown down, but up.

Comparison of the measured kernel with the kernel known to be present shows that the results are quite accurate. The second-degree kernel of the system is known to be

$$g_2(\tau_1, \tau_2) = Kh(\tau_1)\, h(\tau_2),$$

where, from Fig. XIII-7, K is −1.1; and from Fig. XIII-6, h(t) is given.

The period of one cycle of h(t) is 2.95 msec. The period of one cycle of a measured section of the second-degree kernel is 11 sampling intervals, or 2.96 msec.

The maximum height of the known kernel is

$$g_2(\tau_1, \tau_2)\Big|_{max} = K\Big[\, h(t)\big|_{max}\Big]^2 = -1.1 \cdot (52)^2 = -2970.$$

This value agrees with the measured peak of −3000.

It is worth while to compare the very smooth results of Fig. XIII-10, as measured

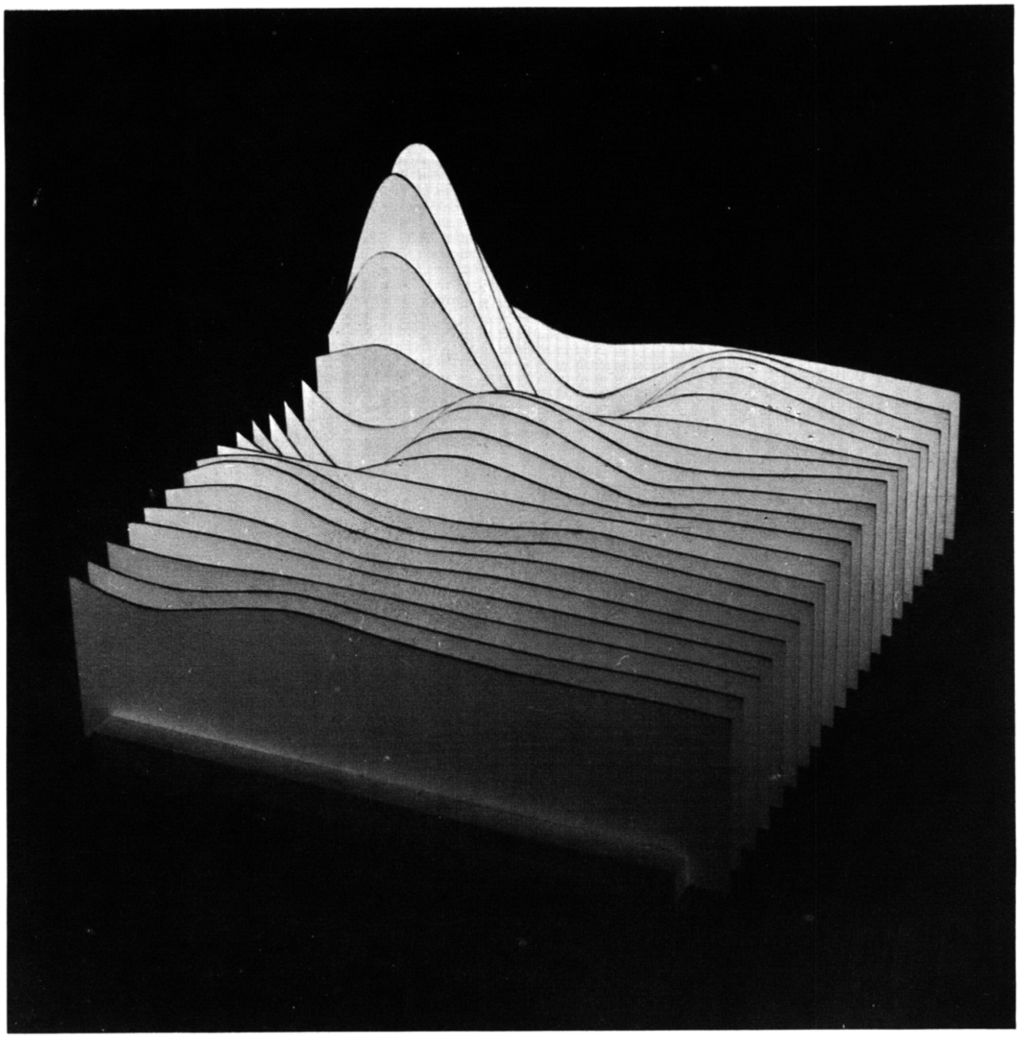


Fig. XIII-9. Three-dimensional model of the measured second-degree kernel.

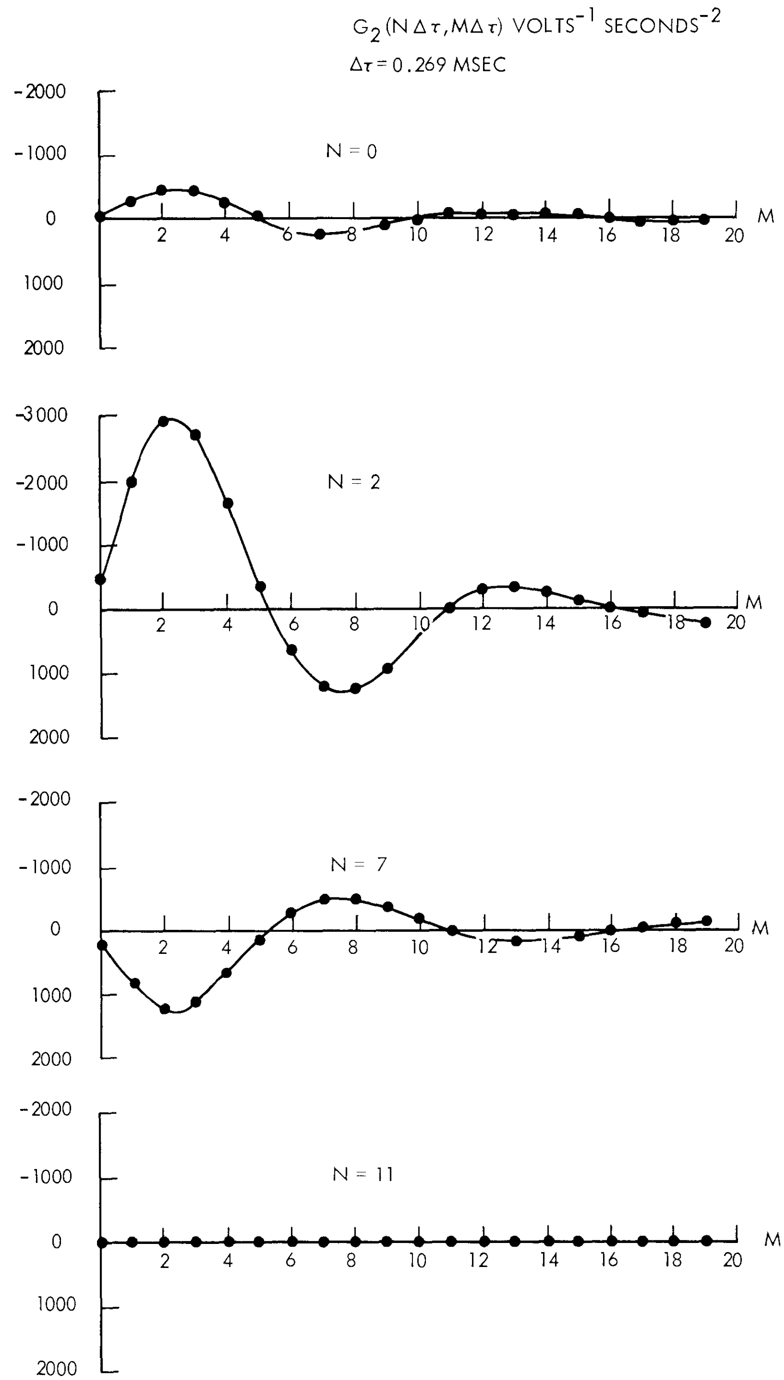


Fig. XIII-10. Second-degree kernel of the nonlinear system as measured by crosscorrelation. (Each point is an average of 29,800 sample products.)

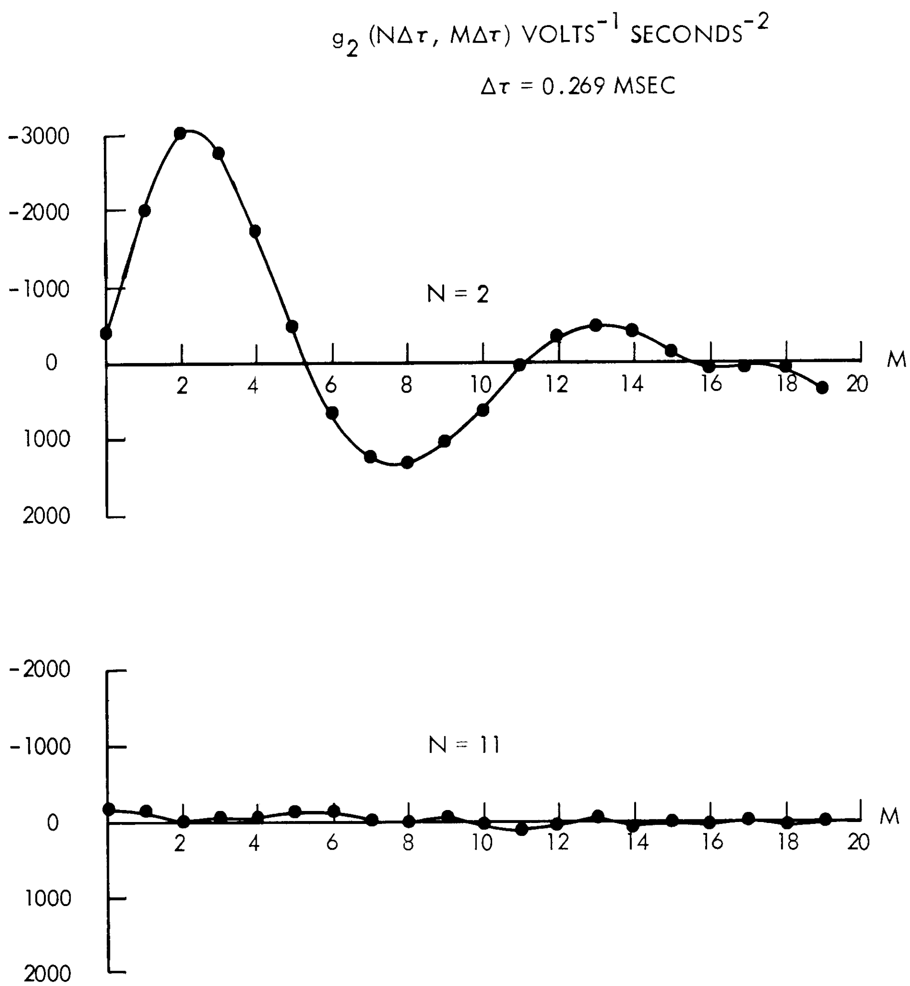


Fig. XIII-11. Measured kernel showing poor convergence with only 1490 sample products used.

with 30,000 independent sample products, with results obtained by using only 1500 independent sample products. A plot of two sections of a 1500-sample product kernel calculation is shown in Fig. XIII-11. In the results for N = 11, there is a noticeable dispersion of the calculated values away from $g_2 = 0$. By accident, N = 11 corresponds to a zero crossing of the known kernel, and therefore provides a convenient line along which to check the statistical error of the kernel measurement. Examining the numerical computer results, we found that the rms value of the 20 points measured for N = 11 is 85 units. Expressed as a percentage of the maximum kernel value of 3000, this yields a statistical error of 2.8 per cent.

The same examination is performed on the results given in the kernel measurement made by using 30,000 sample pairs. The rms deviation is 19 units. This is a statistical error of only 0.6 per cent. In general, the statistical error may be made as small as desired by performing longer crosscorrelations. An average made by using N times as many independent samples as a second average will have a variance that is one N^{th} of the variance of the second average.

4. Conclusions

This work shows that characterization of nonlinear systems by crosscorrelation is experimentally practical. To calculate the second-degree kernel of a system for 20 × 20 points to an accuracy of 0.6 per cent required 15 minutes of IBM 7090 computer time. The same type of calculation to a lesser accuracy of 2.8 per cent required only 45 seconds of computer time.

The question arises: How long will it take to calculate the third-degree, and higher, kernels of an unknown system? A crude estimate can be made by considering the number of elements in each kernel. The second-degree kernel was computed for 20 × 20 points. Because of the kernel symmetry, it was necessary to calculate only one-half of the 400 points. The third-degree kernel would have 20 × 20 × 20 points. In this case, kernel symmetry causes the value of the kernel to be unchanged under any of the 3! or 6 permutations of a given set of arguments. Hence, it is necessary to calculate only one-sixth of the total 20^3 points. The number to be calculated is 20/3 or approximately 7 times the number required for the second-degree kernel. Hence, the calculation of the third-degree kernel would require approximately 7 times the computer time required for the second kernel. This means 5 minutes for 3 per cent accuracy.

Similarly, the calculation of the fourth-degree kernel would require time proportional to $20^4/4!$ which is approximately 33 times greater than that required for the second-degree kernel. This implies 25 minutes of IBM 7090 computer time for 3 per cent accuracy.

Work will continue on the measurement of kernels of several degrees for this system and for other nonlinear systems.

W. S. Widnall

References

1. N. Wiener, Nonlinear Problems in Random Theory (The Technology Press of Massachusetts Institute of Technology, Cambridge, Mass., and John Wiley and Sons, Inc., New York, 1958).

2. Y. W. Lee and M. Schetzen, Measurement of the kernels of a nonlinear system by crosscorrelation, Quarterly Progress Report No. 60, Research Laboratory of Electronics, M.I.T., January 15, 1961, pp. 118-130.

3. W. S. Widnall, Measurement of a Second Degree Wiener Kernel in a Nonlinear System by Crosscorrelation, S.M. Thesis, Department of Electrical Engineering, M.I.T., January 15, 1962.